3. Stochastic Elements

3.1. Introduction and Summary

This chapter contains the contents of the three Letter Reports that we produced in support of Subtask 2, "Stochastic Elements." This subtask is concerned with a technical issue, namely to ensure that the diversities of demographic and economic outcomes at baseline are preserved in the projections. We identified areas in model components of both The Urban Institute/Brookings Institution and RAND where special care was required to preserve such diversities. Our recommendations have subsequently been incorporated in projections of The Urban Institute/Brookings Institution and RAND. As a result, this chapter is no longer necessary to understand MINT. Indeed, the reader may wish to skip to Chapter 4 (page 71). Below we largely replicate reports that were produced under this subtask, without editing the contents in any substantively important way. They reflect an evolving discussion of the relevant issues.

Before replicating the reports, we briefly summarize the issue and our recommendations. As detailed below, variation across individuals in the outcomes of interest is created by many factors. The "stochastic elements" referred to in the Statement of Work are economic and fiscal variables that affect the income and demographic outcomes of interest. However, residual variation is at least of equal importance to ensure that the distribution of outcomes is preserved in projections.

The first letter report discussed the inclusion of residual variation into projections based on a variety of functional forms. That discussion is replicated below as Section 3.2.4. It is our understanding that The Urban Institute/Brookings Institution have added residual variation to all its projections. Similarly, the RAND demographic projections are all based on random draws from duration distributions and thus incorporate hazard models' implicit residual variation.

The main economic and fiscal variables for which we recommend stochastic variation are (1) returns on equity and bonds (mostly relevant for defined contribution plan balances) and (2) employer match rates of defined contribution pension plan contributions. It is our understanding that The Urban Institute/Brookings Institution have incorporated such variation into their projections.

The second letter report, included as Section 3.3 below, developed three alternatives for incorporating stochastic elements. The first option is to draw once from the distribution; the second option is to project the outcome of interest under multiple (Monte Carlo) draws of the stochastic variables and compute the average; and the third option is to replicate the data many times and draw accordingly many values. We recommended the first option for its simplicity, except in cases where the sample size of the subpopulation of interest is small. Few such sample sizes appear to be small, especially after the MINT contractors agreed to extend the projections to 1992 and 1993 SIPP panels. The Urban Institute/Brookings Institution followed our recommendation and drew single values for the stochastic variables.

The third letter report, included as Section 3.4 below, outlined the ways in which MINT incorporates stochastic variation. Prompted by comments from Christopher Bone, it also developed some theoretical considerations for determining the smallest subpopulation for which MINT is capable of generating reliable distributional consequences. There is no single solution to this issue; it depends on the "Type I Error" (related to significance levels) and "Type II Error" (related to power of the predictions) that one is willing to tolerate. As pointed out by John Rust, however, it is always possible to reduce idiosyncratic noise from single replications by running the model simulations sufficiently many times, until additional replications no longer change the outcome(s) of interest by more than a user-specified tolerance level.

3.2. Specify Stochastic Elements of the Model

3.2.1. Objectives

The overall objective of Task 2 is to ensure that the distributions of the income and demographic outcome variables are preserved in the projections. The goal of this subtask is to identify the economic or fiscal variables that should be stochastic in the model, recommend appropriate distributions for them, and identify logical linkages between stochastic elements needed to maintain internal consistency at the person or family level.

3.2.2. Overview

At this stage of the overall project, many details on model structures and implementation strategies are not yet known. It is therefore not always possible to identify exactly where the simulation model would benefit from the implementation of stochasticity in future values of economic or fiscal variables. In the discussion below, we use our best judgment on the structures of the various submodels based on currently available letter reports and the presentations and discussions during the meeting with the Panel of Experts of June 3, 1998. We make several implicit assumptions and recommend alternative scenarios for cases where the model structure appeared insufficiently well-defined.

Variation across individuals in the outcomes of interest is created by many factors. The focus of this section, in accordance with the RFTOP, is on economic and fiscal variables that affect the income and demographic outcomes of interest. In addition, behavioral differences across individuals, such as in savings rates and in job turnover, may also often be important. We will assume that such behavioral variation is either ignored or accounted for in the modeling strategy. We will return to this issue toward the end of this section and illustrate how even very crude models of individual behavior generate additional variation.

A major source of variation in the forecast is uncertainty about the future values of the variables or assumed parameters used to make the forecast. These values may pertain to regressors, to externally determined parameters that are imposed on the projection model (such as the rate of return on assets), or to other assumptions (such as the assumption that all employers match 50 percent of their employees' DC pension plan contributions). Projections that ignore variation in such future values lose part of the resulting variation in the outcomes of interest.

However, in light of the objective of preserving the current population distributions in the projections, a second source of variation is critically important—perhaps even more important than economic and fiscal variables. No model is likely to fit the observed data perfectly, and projections based on expected values are necessarily subject to (at least) the same error rates. One source of variation in the outcomes around the forecast value is therefore the variability not explained by the underlying model. In linear regression models, such as those used to model income components, such *residual variation* is explicitly part of the model. In qualitative outcome models, such as those used in both income and demographic transitions, residual variation may be explicit or implicit in projected (participation/survival) probabilities.

The RFTOP only refers to variation in future values of covariates or assumed parameters. In the next section, we identify economic or fiscal variables that should be stochastic in the model, recommend appropriate distributions for them, and identify logical linkages needed to maintain internal consistency. Section 3.2.4 develops approaches to adding residual variation to the generic types of models the Urban Institute and RAND are likely to apply. Given that these approaches require access to residuals as predicted from the estimation procedure, and in light of the fact that the RFTOP did not refer to residual variation as part of Task 2, we propose that the Urban Institute and RAND incorporate residual variation in their respective projections. We would appreciate an explicit statement on this issue from the Task Manager.

3.2.3. Economic or Fiscal Stochastic Elements

We now discuss stochastic elements of an economic or fiscal nature. The richness of the submodels suggests many candidates for model elements that vary across individuals or households. In this section, we highlight the factors we deem most important as measured by their potential impact on the distribution of projected outcomes. We discuss these factors by Task, to the extent relevant, so as to facilitate subsequent implementation of an approach to incorporating stochasticity.

Part I, Task 2: Project retirement income from assets and savings

This task is concerned with projecting retirement income from non-pension assets and savings. The two predominant determinants of such income are the individual (or household) saving rate and the rate of return received on assets and savings.

The individual saving rate is a behavioral factor that is part of the Urban Institute model. As outlined in John O'Hare' s letter report on this subtask of May 29, 1998, the model will distinguish various saver types (life-cycle savers, precautionary motive savers, and bequest motive savers). Presumably, the model will incorporate variation in savings rates and/or savings patterns across saver types. The implementation is not yet specified, so it is unclear whether the model allows for variation in saving rates within saver types, as present in the population. Urban' s model development will provide insight in the magnitude of this variation; it may be that incorporating additional stochastic variation does little to account for the distribution of asset income. Since saving rates are behavioral, not economic or fiscal, we will not discuss them in any detail.

The rate of return is an average of returns on various asset categories -- such as stocks, bonds, pass book savings, cash holdings, real estate, own business assets, etc. -- weighted by individuals' portfolio composition shares. Since individuals' portfolios differ, the rate of return varies across individuals. Furthermore, rates of returns on a single asset type may vary over time, generating additional variation.

The Urban Institute model considers both total wealth and homeowner equity. It is not yet clear to us exactly what role the rate of return on total wealth (termed the "rate of appreciation") will play in the model. Will it be estimated as part of the model or imposed from an outside source? It is also not clear yet what role homeowner equity will play in the model. It appears that homeowner equity is part of the total wealth measure, so that the rate of return is an overall rate of return.

The rate of appreciation, *r*, is not subscripted by individual or calendar time in O'Hare's model. The assumption that the rate is constant across individuals and over time, however, is probably overly stringent. *We propose that the projections be subject to rates of return that vary both across individuals and over time.*

We now discuss several issues that arise in determining an appropriate distribution for the rate of return.

- Variation over time in the rate of return on various asset categories may be determined from publications such as Ibbotson Associates (1998) and O'Shaughnessy (1998). See, for example, Table 3.1, taken from Ibbotson Associates (1998).

Table 3.1. Summary Statistics of Annual Returns

	Geometric Mean	Arithmetic Mean	Standard Deviation
Large company stocks	11.0%	13.0%	20.3%
Small company stocks	12.7	17.7	33.9
Long-term corporate bonds	5.7	6.1	8.7
Long-term government Bonds	5.2	5.6	9.2
Intermediate-term government	5.3	5.4	5.7
U.S. Treasury Bills	3.8	3.8	3.2
Inflation	3.1	3.2	4.5

These data are compiled from year-by-year total returns on each of the asset categories. Those year-by-year total returns mask, however, that there is substantial variation in the rate of return of assets *within* each category. For example, the large company stocks data are currently based on the Composite Index of Standard & Poor's 500 largest American companies. However, individual investors rarely hold a portfolio that reflects the relative shares in the S&P 500, so that the rates of return experienced by individual investors vary, even in a single year. The standard deviations presented in Table 3.1 are thus underestimates of actual variation.

Publications such as *The Wall Street Journal* may provide data to compute the standard deviation of the rate of return on individual stocks or bonds. Given that most individuals have some level of diversification in their portfolio holdings, the standard deviation of returns on individual stocks and bonds yield an upper bound on the risks experienced by individuals.

To our knowledge, no data source permits a detailed investigation of portfolio shares on individual investors down to individual assets of all types. We therefore need to make a plausible assumption on the rate of return on total assets within broad asset classes.

- Apart from the issue of pooling of assets into major categories, additional variation arises because individuals differ in their relative portfolio shares. There are several surveys that permit an analysis of asset share holdings. Among them are the SIPP, the Survey of Consumer Finances (SCF), and the PSID. A fruitful approach would be to determine the distribution of individuals' portfolio allocations into broad categories as dictated by the survey(s), possibly conditional on observables such as age. In the MINT projection stage, one would draw from the portfolio allocation distribution and from the distributions of rates of return within asset class to obtain a rate of return for an individual or household.

- A final issue relates to the relevant accumulation period or time horizon. The rates of return affect asset accumulation over a long period of time, both before and after retirement. The compound rate of return has a smaller variance than year-by-year rates of return. It may, however, not be necessary to determine the distribution of compound rates of return over various accumulation periods. An approach whereby draws are applied from year-by-year rates of return will result in compound yields with a smaller variance. If there is a strong autoregressive component to rates of return, it would be missed by this approach. A priori, it is our judgment that the costs of an extensive exploration of this issue outweigh the benefits in terms of preserving the distribution of income from assets and savings.

Part I, Task 3: Project retirement income from pensions

For this task, it is important to distinguish defined benefit (DB) and defined contribution (DC) plans.

An important determinant of the income flows from DB plans is job tenure and thus job turnover. As of today, it is not clear whether and how job transitions will be incorporated in the Urban Institute model. Absence of such a model in favor of some fixed-rule assumption would lose much of the heterogeneity that exists in the population, especially for more recent cohorts who are younger at the last survey date. This behavioral factor is likely to generate much of the variation we observe in receipt of DB pensions benefits.

It should be noted that many retirees receive income from both DB and DC plans and that their entitlement may have accrued at various points over their career. This implies that, for example, workers who at the end of the 1990-1991 SIPP panels are only covered under a DC plan may well switch to a job with DB plan coverage during the projection period. The assumption that workers remain on their current job, or any model that does not allow for workers switching between DB and DC plan coverage, would miss much of the variation in both DB and DC pension receipt.[20]

For DC pensions, three determinants with variation across individuals play a particularly important role.

- Employee contribution rate. Employees vary widely in their contribution rates, mostly within the upper legal limit. This behavioral factor will presumably be part of the Urban Institute model.

- Employer contribution match rate. As discussed by Cori Uccello during the meeting with the Panel of Experts of June 3, 1998, there is variation across employers in their contribution match rate. There appears to be an inverse relationship between the employee contribution rate and the employer match rate. A priori, it is not clear whether variation in employer match rates has substantial implications for variation in income from DC plans. *We recommend that the extent of variation be explored and, if it is substantial, that stochastic variation be added to the employer match rate*. The appropriate distribution may be determined from an analysis of IRS Form 5500 data or from the database that the Employee Benefits Research Institute (EBRI) recently collected. This database contains 401(k) plan information for 23,000 plans with over 2.5 million participants and $75 billion of assets. EBRI may be willing to share some of these data or carry out an analysis for the benefit of this project, as indicated by Jack VanderHei in the meeting of June 3, 1998. Alternatively, GAO (1996) and EBRI/Greenwald (1995) offer guidance on this issue.

- The rate of return on plan assets. The issues here are similar to those raised under Task I-2, retirement income from assets and savings. An important difference is that DC plan assets tend to be more diversified into mutual equity and bond funds than non-pension assets. The rate of return on various asset classes thus

[20] Another reason why the issue is relevant for both income from DB and DC plans lies in potential cash-outs of pension rights upon job separation. As much as 60 percent of accumulated pension plans are currently cashed out upon job change (Yakoboski 1997). The cash-out rate is lower among accounts with high balances, so that only 21 percent of account dollars were cashed out. This remains a very sizeable amount of money with potentially large effects on both pension and non-pension wealth accumulation. Both DB and DC plans may be cashed out. An increasing fraction of DB plans (64 percent in 1993) provide the option of a lump-sum distribution (LSD) upon job separation. Among DC plans, 87 percent provides an LSD option upon job separation. Overall, 72 percent of plan participants were able to take an LSD in 1993, up sharply from 48 percent in 1983 (Scott and Shoven, 1996). Logical consistency with Task I-2 (income from assets and savings) requires that pension cash-outs are accounted for in Task I-3.

corresponds reasonably closely to the rates documented in Ibbotson Associates (1998) and O'Shaughnessy (1998). The literature offers some evidence on 401(k) portfolio allocation. See, for example, Yakoboski and VanDerhei (1996); Papke (1998); and Sundén and Surette (1998). *We recommend that stochastic variation across individuals and time be added to the rate of return on DC plan assets*, and that its distribution may be different from the distribution of rate of return on non-pension assets.

Note that variation in the rate of return may have far-reaching implications for income from private savings accounts which may be created if the Social Security system is (partly) privatized. Depending on the degree of freedom covered workers will have in directing their plan asset investments, potentially large differences in retirement income may result. The analysis of rates of return on DC plans will benefit the choice of distribution for the rate of return on private savings accounts.

Part I, Task 4: Develop predictions of partial retirement earnings

An important stochastic element in the prediction of partial retirement earnings is (partial) labor force participation. The Urban Institute letter report by Caroline Ratcliffe and Lawrence Thompson (29 May 1998) and Ratcliffe's presentation during the meeting with the Panel of Experts indicate that Urban Institute's preferred option includes a separate equation for labor force participation. Stochastic variation in this behavioral variable is introduced by drawing a random number.[21] We agree that this procedure generates sufficient variation and do not see other substantively important sources of stochastic variation.

Part I, Task 5: Project Social Security lifetime earnings

We did not identify stochastic elements that will have a substantial effect on variation in the distribution of Social Security lifetime earnings.

Part I, Task 7: Aging of retirement income and assets

Total income from all sources is the sum of individual components. All issues mentioned above apply, but we do not believe that any additional issues arise.

[21] Since labor force participation is part of the model, its variation falls into the "residual variation" category defined above. In Section 3.2.4 we suggest approaches for incorporating such residual variation.

Part II, Task 1: Demographic projections

Our models of marriage formation, marriage dissolution, and mortality do not impose any parameters from outside the model that may be subject to individual variation. We do not believe that any stochastic elements play a substantial role in determining the distribution of demographic outcomes.

3.2.4. Residual Variation

In making predictions, one tends to estimate the future mean and use it as the prediction. This implies that individuals with identical future characteristics (even if known with certainty) are assigned the same prediction, so that part of the variation in initial values is lost. Although not mentioned in the RFTOP, adding residual variation will be important to preserve the distribution of the projected outcomes. We now turn to the incorporation of residual variation. The discussion centers on types of models; each task or subtask may use one or more of these types of models.

Linear models with known residual distribution

The first approach applies to linear models in which the distribution of the residual is known. One then knows the theoretical distribution of the outcomes conditional on the forecast. A simple, but useful, example is the forecast of income based on a log-linear regression, in which the theoretical distribution resulting from residual variation is normal.

Consider the following simple log-linear model:

$$y_i = b'x_i + u_i$$

where y_i is log-income and u_i is distributed normally. Assuming that all future values of the covariates, denoted x_f, are known with certainty, the consistent projection of log-income is $y_f = \hat{b}'x_f$ and the consistent projection of income (accounting for the variance correction) is $\exp\left\{\hat{b}'x_f + \frac{1}{2}\hat{s}_u^2\right\}$. This projection, however, does not incorporate residual variation. Instead, draw a random number u_f from the normal distribution with mean zero and variance $\hat{s}_u^2$ and project future income as $\exp\left\{\hat{b}'x_f + u_f\right\}$. The resulting projection for individual *i* preserves the distribution in the projected population.

Whether the normality assumption holds for any of the income component models which the Urban Institute will estimate is an empirical question.

Linear models with measurable heteroskedasticity structure

A generalization of the known residual distribution discussed above is a residual distribution with known heteroskedasticity structure. For example, the distribution of wealth holdings is known to widen with age. If the residual structure of such outcomes is partly specified in terms of observables, the projection is a straightforward extension of the procedure described above. For example, suppose the residual in the model above is

$$u_i = h(z_i)v_i,$$

where $h(z_i)$ is a known function of observed covariates (such as age) and v_i is distributed *iid* normally, the projected outcome including residual variation is $\exp\left\{\hat{b}'x_f + h(z_f)v_f\right\}$.

Linear models with unknown residual distribution

While models are often specified and estimated under the assumption of a certain residual distribution, the actual distribution is typically unknown. In such cases, consistent projections may be obtained through the "smearing estimator," which explicitly accounts for potential misspecification of the distribution of the residual (Duan 1997). The method takes draws from the empirical distribution of residuals used in the estimation. In the log-linear example, the mean of $\hat{y}_f = \exp\left\{\hat{b}'x_f + u_f\right\}$ is no longer $\exp\left\{\hat{b}'x_f + \frac{1}{2}\hat{s}_u^2\right\}$. The smearing estimate of the mean is $\hat{y}_f = \frac{1}{N}\sum_{d=1}^{N}\exp\left\{\hat{b}'x_f + u_d\right\}$, where N is a (large) number of draws d from the empirical distribution of the residuals. Multiple random draws from the empirical distribution of residuals may be required to consistently estimate the mean forecast value.

This consistently projected mean, however, does not incorporate residual variation. The simplest method to account for residual variation is to only draw once from the empirical distribution of the residuals, so that the projected outcome is $\hat{y}_f = \exp\left\{\hat{b}'x_f + u_d\right\}$. The resulting projections preserve the original distribution of the outcomes, provided that the projection sample is sufficiently large. The 1931-1960 birth cohorts in the combined 1990 and 1991 SIPP samples should yield such a sufficiently large projection sample.

Tobit models

The residual distribution of Tobit models, including the fixed effect Tobit model proposed for modeling Social Security earnings, is by assumption normal. The projection of future values of the outcome (whether that be earnings or the ratio of earnings to the U.S.-wide average Social Security wage) is given by a straightforward generalization of the procedure outlined for linear models with known residual distribution, above:

$$\max\left[0, \min\left[t_f, \exp\left\{\hat{b}' x_f + u_f\right\}\right]\right],$$

where t_f is the future value of the maximum taxable income (or its ratio to the future U.S.-wide average wage). In other words, the projected income value is truncated from below at zero and from above at the maximum taxable wage, so that the correct fraction of individuals is projected to have zero or maximum earnings.

Probit and ordered probit models

Consider an (ordered) probit model for (partial) labor force participation after retirement, i.e., after the individual starts claiming OASI benefits. The probit index function is given by

$$p_i^* = a' x_i + w_i,$$

where $w_i \sim N(0,1)$ and participation is determined by the value of p_i^* relative to one or more thresholds. In a qualitative model like this, a consistent forecast is a participation rate that may not be encountered in the underlying data. Residual variation may be incorporated by drawing a future value of w_i from the standard normal distribution, say, w_f, and comparing $\hat{a}' x_f + w_f$ to the (ordered) probit threshold(s). The resulting projection will be an actual value in the underlying data.

We believe that the procedure proposed by Caroline Ratcliffe and Lawrence Thompson for labor force participation after retirement (as outlined in their Letter Report of May 29, 1998, and in the presentation on June 3, 1998) follows this approach.

Continuous-time hazard models

In continuous-time hazard models, the analogy of a consistent "mean" projection is the expected failure time, i.e., the expected duration of the spell. The expected failure time is given by

$$\hat{T} = \int_{t=0}^{\infty} t f(t) dt,$$

where $f(t)$ is the probability density distribution of the spell's duration. This density is by definition equal to the product of the hazard and survivor functions, $f(t) = h(t) S(t)$. As in all cases discussed above, the projection of expected durations (until marriage, divorce, widowhood, and death) does not preserve the distribution of such outcomes in the population.

One approach to incorporate residual variation is to compute transition probabilities over short duration intervals, draw a random number that is uniformly distributed over the unit interval, and project a transition on the basis of a comparison between the transition probability and the random number. For example, suppose that the probability of getting married for a person with certain characteristics in the next month is 0.02. We draw a random number between zero and one; if that number is less than 0.02, we project that the person got married in the next month. If not, we compute the transition probability in the following month, draw a new random number, et cetera. This is the procedure we discussed in our proposal.

An alternative approach is to draw a random number k between zero and one and compute the failure time $\hat{T}_f$ as the duration at which the survivor function equals the random number: $S\left(\hat{T}_f\right) = k$. The resulting duration follows the same distribution as the underlying sample distribution. This procedure requires formulation of the inverse of the survivor function; for piecewise-linear log-hazard functions (generalized Gompertz) as proposed for demographic transition models, this inverse has a closed-form solution. We anticipate that this second approach is computationally more efficient than the first. Another, minor, advantage is that it generates exact durations rather than time intervals.

3.3. Develop Approaches to Adding Economic Variability

3.3.1. Objectives

The overall objective of Task 2 is to ensure that the distributions of the income and demographic outcome variables are preserved in the projections. The goal of this subtask is to develop approaches for adding economic variability to the projections at the individual level for economic and fiscal variables that were identified in Task 2-1.

3.3.2. Overview

Uncertainty about future values of variables of interest stems from several sources. The first is *residual variation*: no statistical or economic model, no matter how detailed, is capable of fitting historical data perfectly.[22] There is always an implicit or explicit residual term; that term also applies to predicted values. A second source is uncertainty about the future values of the explanatory variables used to make the forecast. Those variables may be (time-varying) covariates or they may be *stochastic elements* such as the rate of return on assets, employer 401(k) match rates, etc. John Rust, in his letter of June 26, 1998, labels these *predetermined variables.* A third source stems from uncertainty about model parameters. Model parameters need to be estimated, are therefore uncertain, and widen confidence intervals of the forecasts. Rust speaks of *estimation noise*.

The objective of this task is to ensure that the distributions of projected outcomes correspond closely to the distributions of observed outcomes in the SIPP. The distributions need not and probably should not be identical: we would expect wages to grow over time, life expectancy to increase, etc. The RFTOP states that "the diversity of the American population ... should be maintained as the individuals are aged in the projection model." In our interpretation, the objective is thus mostly aimed at preserving the variance of projections.

Given this interpretation, we think that uncertainty about model parameters should not be considered. For example, an overestimate of the growth rate of wages results in higher projected wages for all simulants. It affects the mean projection, with only second-order effects, if any, on the variance. In other words, estimation noise is of limited relevance to diversity and inequality analyses. Indeed, as noted by Rust, "the amount of estimation noise is a third order problem compared to the first order problem of describing the uncertainty in the path of the predetermined variables." This is not to say that it would not be useful to simulate the model many times with different draws from the parameter distribution. Such an exercise would be a form of sensitivity analysis to the values of imprecisely estimated parameters.

---

[22] Except in a specific sample when the number of parameters is increased until zero degrees of freedom. Drawing a new sample would no longer result in a perfect fit.

To preserve the diversity of the population in the projections, we do need to account for residual variation and stochastic elements. The distinction is often a matter of model specification. If a variable explicitly enters a model, its variation causes stochastic variation. If it is omitted, its variation becomes part of residual variation. The distinction may have caused some misunderstanding. For example, our Letter Report on Task 2-1 (replicated here as Section 3.2) stated that we "do not believe that any stochastic elements play a substantial role in determining the distribution of demographic outcomes." Rust took issue with that statement, arguing that "there is considerable uncertainty about things like divorce, remarriage, death of a spouse, and so forth..." We fully agree and have no illusions that our model projects perfectly. What we intended is that all uncertainty is of residual nature; the explanatory covariates in our demographic models are all observed in the data and do not have uncertain future values (e.g., for education, race, and the like there is no source of stochastic variation). In other cases, (e.g., for future marital status transitions) these covariates are explicitly modeled and accounted for.

3.3.3. Residual Variation

As indicated in our Letter Report on Task 2-1 (Section 3.2), we propose that the Urban Institute and RAND both incorporate residual variation in their projections. Doing so is a very natural element of the projection procedure. Furthermore, if RAND were to add residual variation to Urban's projections, Urban staff will need to spend considerable time documenting every estimation and projection program, all data will need to be transferred, and all projection programs will need to be re-run. In other words, the fixed costs would far exceed the very low costs of adding residual variation. It would also delay the moment at which the Near Term Model may be applied to policy simulations, because projections without residual variation have more limited value for studying distributional consequences and inequality.

Section 3.2.4 specified how residual variation may be added. Note that no Monte Carlo simulations are involved; only a single random number needs to be added to every projected outcome. For binary choice models, the projection technique naturally and almost inevitably incorporates residual variation. For example, the decision to engage in part-time work after retirement is projected by adding a normally distributed random number to the probit index function to project whether an individual retiree will work. For linear and Tobit models, adding a random draw from the residual distribution is trivial. For hazard models, the issue is more complicated, though not more complicated than the computation of an expected duration. RAND will incorporate residual variation in its demographic hazard models.

Naturally, we will provide any assistance necessary to clarify the procedure.

3.3.4. Stochastic Variation

We propose three options for incorporating stochastic variation. All three assume that stochastic variables and their distributions have been identified, as discussed in Section 3.2. The first option is to replace the constant value of a stochastic element by a draw for every individual (and every time period or every employer or every marriage, as appropriate); the second is to conduct Monte Carlo simulations; and the third is based on a single large replication.

The RFTOP also mentions the Ibbotson Associates technique of estimating final values of variables such as savings and assets (Ibbotson Associates, 1998, Chapter 9). Our interpretation of that technique is that Ibbotson draws from a known statistical distribution (the book mentions the log-normal distribution as appropriate for asset return relatives) rather than from an empirical distribution. As such, we think that this establishes the distribution from which stochastic elements are to be drawn. It therefore relates to a preliminary step of incorporating stochastic variation. We will apply either technique as appropriate for the stochastic element under consideration.

Option 1: A Single Set of Draws

Since none of the income projection methods has been finalized, it is somewhat difficult to illustrate techniques through examples. We will make some assumptions in the following discussion; some of those may not apply.

In the first option, we replace a constant variable—which should really be subject to variability—by a single draw or set of draws. For example, the model that projects income from assets contains a parameter representing the rate of return on assets. The baseline model may set that parameter to a constant value for all individuals and all time periods. In the first option, we draw rates of return for every individual and every year from the distribution of annual rates of return (as specified in Section 3.2) and substitute it for the constant model parameter. The resulting distribution of incomes from assets will have a variance that is greater than the variance under the assumption of a constant rate of return (provided that the distribution of rates of return is not in some very systematic way negatively correlated with asset holdings).

If the simulation sample were infinitely large, this technique would result in the appropriate distributions of the outcomes of interest. The combined 1990 and 1991 SIPP simulation sample has roughly 50,000 individuals, which may not be sufficiently large to support appropriate distributions on certain rare subpopulations. The next options remedy this.

Option 2: Monte Carlo Simulations

A natural way to implement stochastic variation is through Monte Carlo simulations. Consider again projections of income from assets that may be strongly affected by the rate of return on assets. The projection program would be run many times, say, 100 times. The rate of return on assets would not be constant, but vary across individuals and across time periods. This is achieved by drawing a rate of return for every individual and every time period. As the projection model is run 100 times, 100 sets of rates of returns will be drawn, resulting in 100 different projections for every individual.

The results should be aggregated first over individuals and second over Monte Carlo iterations. Suppose one is interested in income inequality as measured by a Gini coefficient. Each Monte Carlo iteration results in income projections for all 50,000 individuals in the sample. These should first be aggregated into a Gini coefficient, resulting in 100 Gini coefficients. In the second step, the mean Gini coefficient and its standard deviation may be computed.

Option 3: One Large Replication

A third alternative essentially combines the first two options. One may take the simulation sample of 50,000 individuals and duplicate them a large number of times, say, 100 times. This results in a simulation sample of 5,000,000 individuals. Assign all individuals (sets of) draws from the distribution of the stochastic element and run the projection program. The resulting projections support analyses of distributional consequences even for rare subpopulations, because each member of such population is represented 100 times.

One virtue of this approach is its simplicity. Duplicating observations is trivial, and the projection program needs to be run only once. A disadvantage is that it does not yield an estimate of the variance of the summary statistic of interest. The Gini coefficient that is based on 5,000,000 individuals is equal to the mean coefficient obtained from Monte Carlo simulations (up to nonlinearities), but no standard deviation is obtained, i.e., the technique is not so rich in the information that it generates. Another disadvantage is that the technique may require substantial amounts of disk storage space. A minor advantage, finally, is that the resulting projection data support the computation of alternative summary statistics. A Monte Carlo simulation would need to be re-run, since the summary statistic would need to be computed for every iteration.

Table 3.2 summarizes the advantages and disadvantages associated with the three options. Two asterisks imply a higher level of attractiveness.

Table 3.2. A Comparison of the Three Options

	Ability to analyze rare subpopulations	Information richness	Simplicity	CPU time	Disk storage	Ability to generate alternative summary measures
1	*	*	**	**	**	**
2	**	**	*	*	**	*
3	**	*	**	*	*	**

We recommend using option 1 except in applications where rare subpopulations are being studied or where there is some other compelling reason to gain the efficiency of multiple replications. The organizational and coordination costs of performing multiple replications for Monte Carlo or large replication may usually out-weigh the benefits. If multiple replications are performed, retaining the replication number either option 2 or option 3 can be performed ex post.

3.4. Implement Technique for Adding Economic Variability

3.4.1. Objectives

The overall objective of Task 2 is to ensure that the distributions of the income and demographic outcome variables are preserved in the projections. Task 2-1 identified economic and fiscal variables that should be subject to stochasticity. Task 2-2 developed techniques for implementing such stochasticity into the projection models. The goal of Task 2-3 is to implement the technique as selected by the Task Manager for adding economic variability to the projections at the individual level.

3.4.2. Stochastic Variables

Section 3.2 identified the following economic variables as important for preserving the distribution of projected income flows and demographic states.

- The rate of return on assets and savings, for projecting retirement income from assets and savings (Part I, Task 2). We recommended that the projections be subject to rates of return that vary both across individuals and over time.
- The employer contribution match rate, for projecting retirement income from defined contribution pensions (Part I, Task 3). We recommended that the projections be subject to employer contribution match rates that are an inverse function of the employee contribution rate, and vary across individuals.
- The rate of return on DC plan account balances, for projecting retirement income from defined contribution pensions (Part I, Task 3). We recommended that the projections be subject to rates of return that vary both across individuals and over time.

At the time we identified these stochastic variables, no decision had been made on the structure of The Urban Institute’ s model underlying retirement income from assets and savings (Part I, Task 2). The SSA Task Manager opted for a model in which age-wealth profiles are estimated and projected and in which the rate of return on assets and savings no longer explicitly enters. Implicitly, of course, the rate of return still plays a central role, but year-on-year and cross-individual variation in the rate of return is now part of the residual structure.

Our reports emphasized that, in addition to stochasticity of economic and fiscal variables, the residual structures of the various component models in MINT play an important role in maintaining the distribution of income and demographic outcome variables of future projections. We further discuss residual variation in the next section.

3.4.3. Techniques

Sections 3.2 and 3.3 distinguished residual variation from stochastic variation. Both play an important role in maintaining distributions of outcomes in projections. For practical purposes, the distinction is largely a function of model structure and specification. If a certain parameter enters the model and is assumed fixed (but really varies across individuals, over time, or otherwise), projections will exhibit too little variation. If a parameter is omitted from the model, its variation is absorbed by the residual term; projections that set future residual terms to their mean value (usually zero) will therefore again exhibit too little variation.

We recommended that both The Urban Institute and RAND include residual variation into their projections. This issue was further discussed in a meeting between Howard Iams, Lee Cohen, Eric Toder, Gary Burtless, and Stan Panis on 20 July 1998. Iams subsequently directed The Urban Institute and RAND to add single draws from residual distributions to each party' s projections. Details on appropriate procedures for various types of models are provided in Section 3.2.

In Task 2-2 we developed three techniques for incorporating stochastic variation: (1) substitute a single draw from the stochastic variable' s distribution for its mean value; (2) repeatedly substitute draws from the stochastic variable' s distribution for its mean value, and compute aggregates of the projected outcomes of interest (known as Monte Carlo simulations); and (3) replicate the simulation sample and draw once from the stochastic variable' s distribution, essentially leading to similar results as the second option. We recommended the first option: draw once from the stochastic variable' s distribution and use it, instead of a constant value, to project the outcome of interest.

As stated above, projections of income from assets and savings no longer involve a rate of return. The two remaining stochastic elements that we identified as important, the employer match rate of DC plan contributions and the rate of return on DC plan balances, both apply to The Urban Institute' s Task 3 (retirement income from pensions). The SSA Task Manager directed The Urban Institute to draw single values from the empirical distributions of employer DC plan contribution match rates and rates of returns on stocks and bonds and to substitute these draws into the projection model.[23] This directive is consistent with our recommendation.

---

[23] The distribution of employer match rates is based on 1995 Survey of Consumer Finances (SCF) data; the distribution of rates of return on the stock portion of DC plan balances is based on S&P 500 returns from 1952-1994; the distribution of rates of return on the fixed-income portion of DC plan balances is based on T-Bill rates over the same period. The rates of return are assumed to be normally distributed (in slight deviation from Ibbotson, 1998, but consistent with Cohen, 1998) with means equal to variances. At the time of this report, the exact parameters have not yet been decided upon. Cohen (1998) found 1952-1994 average real rates of return for the S&P 500 stock index of 6.98 percent and 1.17 for T-Bills. See Letter Report for Task 3-2 of The Urban Institute.

The Optimal Number of Draws

Christopher Bone, in his memorandum to Howard Iams of July 24, 1998, raised an important and intellectually challenging issue. Section 3.3 asserted that a single set of draws to replace constant variables is appropriate except where rare subpopulations are being studied or there is some other compelling reason to gain the efficiency of multiple replications. Bone pointed out that this does not provide those using the model with any guidance on what constitutes a "rare" subpopulation. Furthermore, cumulative mortality increases the number of relatively rare subpopulations: a subpopulation may not be "rare" at age 62, but may be "rare" in 2020.

Note that Bone's critique applies to residual variation as well as to stochastic variation. The Urban Institute and RAND both include a single draw from the residual distribution into the projections. Ideally, to get the full distribution, very many draws should be included.[24]

Bone went on to request the development of a method for classifying the number of replications needed to ensure sufficient variation in the population.

The issue is important not only for its relevance to stochastic variation, but also because it generalizes to MINT as a whole: "*What is the smallest subpopulation for which MINT is capable of generating reliable distributional consequences?*" Obviously, there is no hard and uniformly applicable answer to this question. A core feature of any statistical prediction model is that its predictions have smaller confidence intervals and larger power the larger the subpopulation of interest is. Rather than solve the issue in general terms, we now discuss the factors relevant to the subpopulation size issue.

For discussion purposes, consider the following thought experiment. A MINT user wants to figure out whether a certain policy measure (or combination of policy measures) results in an increased poverty rate by 2020 among, say, elderly African-American women living alone. Assume for now that we follow the best possible procedure (within more fundamental limits of the MINT component models), namely to draw very many times for both stochastic variables and residuals. The user runs MINT twice: once under baseline assumptions (no change in policy) and once incorporating the package of policy measures. Suppose, for argument's sake, that the baseline poverty rate in 2020 among the subpopulation of interest is projected to be 20 percent ($\hat{p}_1 = 0.20$) and the post-policy rate 23 percent ($\hat{p}_2 = 0.23$). These rates have distributions, because the parameters underlying MINT are not known with certainty, but have been estimated. We are interested in the change in poverty rate ($d = p_2 - p_1$), for which we have a point estimate of 3 percent ($\hat{d} = 0.03$). Its standard

[24] The issue is actually less relevant for stochastic variation of rates of return than for other variation. For projecting the value of a DC plan upon retirement, the relevant rate of return is the rate as it compounds over all accumulation years. Since a rate is drawn for every year anew, the compounded rate is very stable.

deviation may, in principle, be estimated.[25] The standard deviation depends inversely on the size of the subpopulation.

To determine required subpopulation sample sizes, two types of error are relevant.

The first type of error, known in the statistical literature as "Type I error," arises when the null hypothesis is true but rejected by the simulation. Suppose the policy measure under simulation does not affect the poverty rate of the subpopulation of interest, i.e., the null hypothesis that $d = 0$ is true. Whether the point estimate, $\hat{d} = 0.03$, is large enough to reject the null hypothesis depends on the standard deviation of $\hat{d}$ and the chosen significance level. In practice, a probability of a Type I error of 5 percent tends to be considered acceptable. That 5 percent probability ("significance level") corresponds to 1.96 standard deviations, so if 0.03 exceeds 1.96 times the standard deviation of $\hat{d}$, the MINT user would conclude that there is a significant increase in the poverty rate, and thus make a Type I error.

The second type of error, known as "Type II error," arises when the alternative hypothesis is true but rejected. This type of error is directly related to the "power" of a test. Power is defined as one minus the probability of a Type II error. To assess the power of a test, the MINT user must take a stance on an alternative hypothesis of interest. For example, the user may decide that increases in the poverty rate of less than 4 percentage points are of no great concern, but that the model must be capable of detecting changes of 4 percentage points or more. This corresponds to an alternative hypothesis that $d = 0.04$. Would the user reject that hypothesis? The answer depends again on the standard deviation of $\hat{d}$ and the probability of a Type II error that the user is willing to accept. In practice, a probability of a Type II error of 20 percent tends to be considered acceptable; put differently, users tend to require a power of at least 80 percent. This 20 percent probability corresponds to 0.84 standard deviations, so if 0.04-0.03=0.01 exceeds 0.84 times the standard deviation of $\hat{d}$, the MINT user would reject the hypothesis that the poverty rate increased by 4 percentage points, and thus make a Type II error.

Note that both types of errors depend on the size of the subpopulation of interest. Larger sample sizes reduce the probabilities of making Type I and Type II errors. In

[25] Note that d is a function of model parameters, q, characteristics of the subpopulation, X, and policy parameters. Policy parameters and respondent characteristics are assumed to be known with certainty, so we may approximate the variance of d by:

$$\hat{s}_d^2 \approx \frac{\partial d}{\partial q'} \hat{\Sigma}_{qq} \frac{\partial d}{\partial q},$$

where $\partial d / \partial q$ is the (numerically computed) first derivative of the increase in poverty rate with respect to MINT parameters and $\hat{\Sigma}_{qq}$ is the covariance matrix of the model parameter estimates. The latter, $\hat{\Sigma}_{qq}$, is not typically reported in model estimates, but is accessible in most software packages. It is block-diagonal in the MINT model, because model modules have been estimated separately, under the assumption of independence. The computation will be hugely difficult and time-consuming, but is possible.

principle, it is possible to compute the sample sizes that are required to achieve a significance level and power that are selected as sufficiently high.

The above assumed that we conducted Monte Carlo simulations with infinitely many draws for both residual and stochastic variation. As the number of draws goes down to a finite number, and perhaps as low as one (such as is currently the case), the standard deviation of $\hat{d}$ becomes subject to uncertainty about residual and stochastic values, in addition to the uncertainty from imprecisely estimated parameters. Its computation becomes extraordinarily complex.

Approaching the issue from the opposite direction, we investigated the number of respondents in the simulation sample that represent subpopulations of actual policy interest. The combined 1990, 1991, 1992, and 1993 SIPP panels contain 65,369 full panel respondents born in 1931-1960. Of these, 17,998 are projected to become deceased before the year 2020, and 47,371 (72.5 percent) are projected to survive. Table 3.3 shows the number of respondents among the 47,371 survivors that are part of subpopulations of interest. To indicate the political clout of these subpopulations, the table also shows population-weighted figures, i.e., the numbers of Americans that are represented by the SIPP respondents.[26]

Table 3.3. Selected subpopulations in the year 2020

	Simulation sample size	Weighted to U.S. population
African American women living alone age 65-89	1,858	2,912,285
Native American women living alone age 65-89	108	164,778
Hispanic women living alone age 65-89	1,172	1,594,340
African American women living alone age 75-89	747	1,095,386
Native American women living alone age 75-89	48	76,014
Hispanic women living alone age 75-89	504	632,902
Divorced women who have been married 10+ years	7,962	11,920,746
Individuals age 85-89	2,213	3,174,837

Note that the numbers of SIPP respondents that are members of these subpopulations and survive through the year 2020 are fairly high. For most subpopulations we have more than 500 sample members. At issue is the number that is required to determine a distribution, or, of keener political interest, the fraction of individuals below the poverty line. Without working through the formal model outlined above, it is impossible to know such minimum sample sizes. Intuitively, one may feel comfortable that 50-100 observations are sufficient to pin down changes in a poverty rate, especially if poverty is a relatively common occurrence, such as it is among elderly Native American women.

[26] Note that the oldest person, born in 1931, is only 89 years old as of January 1, 2020. As much as 32 percent of the individuals born in 1931-60 are expected to live past their 90-th birthdays, so projections of the oldest-old may gain precision from incorporating pre-1931 cohorts in the projections.